

# Improvement of the Bayesian neural network to study the photoneutron yield cross sections\*

Yong-Yi Li,<sup>1</sup> Fan Zhang (https://orcid.org/0000-0003-3642-6639),<sup>2,†</sup> and Jun Su (https://orcid.org/0000-0002-8518-6443)<sup>3,‡</sup>

<sup>1</sup>Department of Computing, Changzhi University, Changzhi 046011, China

<sup>2</sup>Department of Physics, Changzhi University, Changzhi 046011, China

<sup>3</sup>Sino-French Institute of Nuclear Engineering and Technology, Sun Yat-sen University, Zhuhai 519082, China

This work is an attempt to improve the Bayesian neural network (BNN) for studying photoneutron yield cross sections as a function of the charge number  $Z$ , mass number  $A$ , and incident energy  $\varepsilon$ . The BNN was improved in terms of three aspects: numerical parameters, input layer, and network structure. First, by minimizing the deviations between the predictions and data, the numerical parameters, including the hidden layer number, hidden node number, and activation function, were selected. It was found that the BNN with three hidden layers, 10 hidden nodes, and sigmoid activation function provided the smallest deviations. Second, based on known knowledge, such as the isospin dependence and shape effect, the optimal ground-state properties were selected as input neurons. Third, the Lorentzian function was applied to map the hidden nodes to the output cross sections, and the empirical formula of the Lorentzian parameters was applied to link some of the input nodes to the output cross sections. It was found that the last two aspects improved the predictions and avoided overfitting, especially for the axially deformed nucleus.

Keywords: Bayesian neural network, Photoneutron cross sections

## I. INTRODUCTION

Neural networks are powerful tools for making predictions after training using data and have made possible exciting achievements in nuclear physics in the past few years [1–6]. The earliest work on neural networks in nuclear physics dates back to 1993, when the phenomenological approach to many-body systems based on multilayer feedforward neural networks was introduced to learn the systematics of atomic masses and nuclear spins and parities [7]. Thereafter, various types of neural networks have been applied to study nuclear mass systematics [8, 9],  $\beta$ -decay systematics [10], and binding energy [11].

Recently, with the help of physical ideas, neural networks have been improved and their potential capability has been realized. The known physics were explicitly embedded based on the Bayesian neural network (BNN), which results in a novel method for accurately predicting  $\beta$ -decay half-lives [12]. The input data are preprocessed, correlations among the input data are included, and the problem of multiple solutions can be reduced, yielding more stable extrapolated results [13]. The combination of the three-parameter formula and the BNN results in a novel approach for describing the charge radii of the nuclei [14]. In addition, the physical law is revealed by neural networks. For example, a feedforward neural network model was trained to calculate nuclear charge radii, and the correlation between the symmetry energy and charge radii of Ca isotopes was suggested [15]. The convolutional neural network algorithm was applied to determine the impact parameters in heavy ion collisions using constrained molecular-dynamics model simulations [16]. To date, neural networks have been widely applied for sig-

nal identification [17–19], data restoration [20, 21], and regression analysis [5] and have been broadly used in nuclear physics.

In studies where the nuclear masses and nuclear charge radii were studied, Utama et al. claimed that the physics in the initial prediction can be included using physics-motivated models and that the BNN can be used to fine-tune these models by modeling the residuals [22, 23]. Based on this residual-approach, the predictions of several physics-motivated models in nuclear physics have been improved using neural networks. For example, the BNN approach was employed to improve the nuclear mass predictions of various physics models [24] and the fission yield prediction using the TALYS model [2]. To study the isotopic cross sections in proton-induced spallation reactions, the BNN predictions obtained by learning the experimental data and the residuals of the SPACS parametrization were compared [25]. It was shown that the latter are better, which indicates the significance of the robust physics-motivated model when using a neural network.

A neural network is a type of numerical algorithm. In cases where a physics-motivated model is not available, attempts have been made to provide physics guidance in neural networks. A multitask neural network was applied to learn the experimental data of the giant dipole resonance parameters directly, and the predictions were better than those calculated by the Goldhaber—Teller model [26]. In a study for determining the impact parameters of heavy-ion collisions using convolutional neural networks, no initial predictions were made [16]. In addition to the physics-motivated model, physics guidance in neural networks has also been provided from the input layer [27] or by an empirical formula [14]. In this study, an attempt is made to improve the BNN for studying photoneutron yield cross sections, where both the improvements from the input layer and by the empirical formula are considered.

Photonuclear reactions were first observed more than 60 years ago [28]. Their cross section data are important to analyze the radiation transport, study the nuclear waste transmu-

\* This work was supported by the National Natural Science Foundation of China (Nos. 11905018 and 11875328).

<sup>†</sup> zhangfan@mail.bnu.edu.cn

<sup>‡</sup> sujun3@mail.sysu.edu.cn

tation [29], and calculate the nucleosynthesis [30, 31]. The underlying mechanisms in photonuclear reactions are significant in fundamental nuclear physics [32, 33]. More than 27000 data of  $\sigma_{xn}$  were collected in the EXFOR database [34] for nuclei from  ${}^6\text{Li}$  to  ${}^{239}\text{Pu}$  at incident energies above the neutron separation energy. New facilities have been developed to measure additional data [35, 36]. A large dataset makes machine learning possible and advisable.

This study focuses on the improvement of the BNN to study photoneutron yield cross sections. The remainder of this paper is organized as follows. In Sect. II, the model is described. In Sect. III, both the results and discussions are presented. Finally, a summary is presented in Sect. IV.

## II. MODEL

The fundamentals of the BNN approach were established in the last century [37]. It is now a commonly used method for pattern recognition and numerical regression. The latter is the case in the present study. In the normal BNN algorithm, a neural network with hidden layers is built for mapping from the input layer  $X$  to the output layer  $Y$ .

$$Y(X, \theta) = a + \sum_{j=1}^{N_h} b_j f \left( c_j + \sum_{i=1}^l d_{ji} X_i \right), \quad (1)$$

where  $\theta = \{a, b_j, c_j, d_{ji}\}$  are parameters in the neural network,  $\{a, c_j\}$  are biases, and  $\{b_j, d_{ji}\}$  are weights. The activation function  $f$  can be set as a sigmoid, tanh, or softplus type.

$$\begin{aligned} \text{sigmoid}(t) &= \frac{1}{1 + \exp(t)}, \\ \tanh(t) &= \frac{\exp(t) - \exp(-t)}{\exp(t) + \exp(-t)}, \\ \text{softplus}(t) &= \log(1 + \exp(t)). \end{aligned} \quad (2)$$

With the activation function, nonlinearity between the input and output is realized.

The parameters in Eq. (1) are determined by learning dataset  $D$ , which displays the outputs  $Y$  for given inputs  $X$ ,

$$D = \{X^{(n)}, Y^{(n)}\}_{n=1}^{N_d}, \quad (3)$$

where  $N_d$  is the sample size of the available data. The Bayesian theorem, which deduces the posterior knowledge from the prior case using the likelihood function, is applied to solve this regression problem. Specifically, the prior distribution  $P(\theta)$  of the parameters  $\theta$  is assumed, and its posterior distribution  $P(\theta|D)$  for a given dataset  $D$  is expressed as

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{\int P(D|\theta)P(\theta)d\theta}, \quad (4)$$

where  $P(D|\theta)$  is the likelihood of  $D$  given  $\theta$ .

With the posterior distribution of the parameters  $\theta$ , the expected value of the output  $\hat{Y}^*$  for the known inputs  $X^*$  is expressed as the integration

$$\hat{Y}^* = \int Y(X^*, \theta) P(\theta|D) d\theta, \quad (5)$$

Monte Carlo techniques are applied to calculate the above integration.

$$\begin{aligned} \hat{Y}^* &= MCS_{P(\theta|D)} [Y(X^*, \theta)] \\ &\approx \frac{1}{N_s} \sum_k^{N_s} Y(X^*, \theta^{(k)}), \end{aligned} \quad (6)$$

where  $MCS_{P(\theta|D)}$  denotes the Monte Carlo sampling drawn from the posterior distributions  $P(\theta|D)$  and  $\theta^{(k)}$  ( $k = 1, 2, \dots, N_s$ ) is the  $k$ -th sample drawn from the posterior distribution  $P(\theta|D)$  with total sample number  $N_s$ . The accuracy of Eq. (6) to evaluate Eq. (5) is expressed as

$$Y^* = \hat{Y}^* \pm 1.96 \frac{S}{\sqrt{N_s}}, \quad (7)$$

where  $S$  is the standard deviation of the samples  $Y(X^*, \theta^{(k)})$ . The difference between the integration and Monte Carlo calculations can be reduced by increasing the number of samples  $N_s$ .

The analytic computation of the posterior distribution  $P(\theta|D)$  is intractable owing to the high dimensionality of the parameters. In the BNN approach, variation inference is applied to obtain an approximation of  $P(\theta|D)$ . The variation inference attempts to find a  $\kappa$  such that  $q(\theta|\kappa)$  is the minimum distance from  $P(\theta|D)$  measured by the (KL) divergence:

$$\begin{aligned} \theta &= \arg \min \text{KL} [q(\theta|\kappa) || P(\theta|D)] \\ &= \arg \min E_{q(\theta|\kappa)} \left[ \ln \frac{q(\theta|\kappa)}{P(\theta|D)} \right] \\ &= \arg \min E_{q(\theta|\kappa)} \left[ \ln \frac{q(\theta|\kappa)P(D)}{P(D|\theta)P(\theta)} \right] \\ &= \arg \min \sum_k \left[ \ln q(\theta^{(k)}|\kappa) - \ln P(\theta^{(k)}) \right. \\ &\quad \left. - \ln P(D|\theta^{(k)}) \right]. \end{aligned} \quad (8)$$

The traditional method of solving the regression problem in physics is based on an empirical formula with parameters, where an appropriate formula can avoid both misconvergence and overfitting. Similarly, when applying a BNN in physics, it is important to select the appropriate input nodes for a specific output. In this study, the output is the photoneutron yield cross section  $\sigma_{xn}$ . More than 27000 data of  $\sigma_{xn}$  were collected in the EXFOR [34] for nuclei from  ${}^6\text{Li}$  to  ${}^{239}\text{Pu}$  at incident energies above the neutron separation energy. The minimum input nodes of the BNN to study  $\sigma_{xn}$  are the charge number  $Z$  of the target, mass number  $A$  of the target, and energy  $\varepsilon$  of the incident  $\gamma$  particle. The BNN in this case is abbreviated as BNN-ZAE and illustrated in Fig. 1(a).

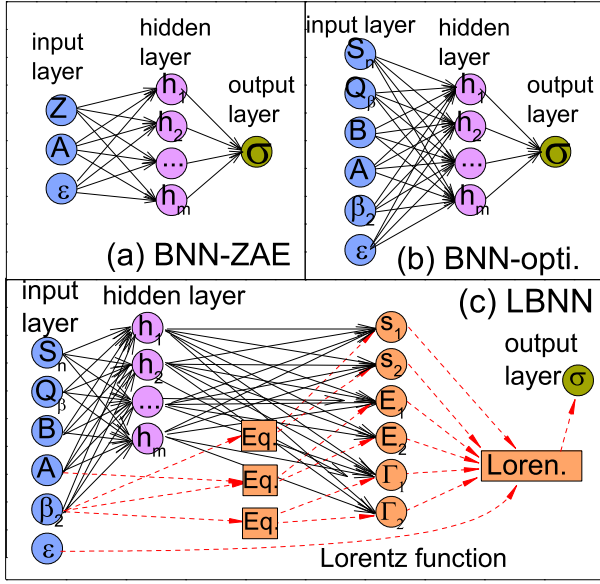


Fig. 1. (Color online) (a) Bayesian neural network BNN-ZAE with input nodes of charge number  $Z$ , mass number  $A$ , and incident energy  $\epsilon$ . (b) Bayesian neural network BNN-OPT with optimal input nodes including mass number  $A$ , incident energy  $\epsilon$ , and other ground-state properties. (c) Lorentzian function-based Bayesian neural network LBNN where the Lorentzian shape of the photoneutron cross sections and the empirical formula of the Lorentzian parameters are considered.

The BNN is a numerical algorithm. In physics, the BNN is used to learn the residuals of the physics-motivated model and then fine-tune the model. Thus, the main physics information is included in the initial prediction by the physics-motivated model. In our previous work [38], we illustrated a new method to provide physical guidance in the BNN from the input layer without initial prediction by the physics-motivated model. This method is applied in this work to select the optimal ground-state properties as neurons of the input layer in the BNN to predict the photoneutron yield cross sections  $\sigma_{xn}$ . Details of the method are provided in Ref. [38]. In brief, various combinations of ground-state properties are considered as the input nodes in the BNN, and the optimal combination is selected according to the smallest deviation between the data and the prediction. The optimal input nodes are as follows:

$$X = \{S_n, Q_\beta, B, A, \beta_2, \epsilon\}, \quad (9)$$

where  $S_n$  is the neutron separation energy,  $Q_\beta$  is the  $\beta^-$  decay energy,  $B$  is the binding energy per nucleon,  $A$  is the mass number,  $\beta_2$  is the quadrupole deformation parameter, and  $\epsilon$  is the incident photon energy. In this case, the BNN is abbreviated as BNN-OPT and illustrated in Fig. 1(b). It is worth noting that the charge number  $Z$  is not input in the BNN-OPT model. Only the experimental data of the photoneutron yield cross section for stable nuclei are available. For stable nuclei, the correlation between the charge number  $Z$  and the mass number  $A$  is very strong (see Fig. 1 in Ref. [38]). After mass number  $A$  has been used, the additional  $Z$  is not conducive to

reducing the deviation between the data and the prediction by the BNN-OPT model.

We will further prove that the BNN can be improved by considering known knowledge of physics. The photoneutron yield cross sections  $\sigma_{xn}$  as a function of incident energy can be characterized by a Lorentzian shape with two components. The Lorentzian parameters are peak energy  $E_i$ , width  $\Gamma_i$ , and strength  $s_i$ .

$$\sigma_{xn} = \sum_{i=1,2} \frac{2}{\pi} \frac{\sigma_{\text{TRK}} s_i \epsilon^2 \Gamma_i}{[\epsilon^2 - E_i^2]^2 + (\epsilon \Gamma_i)^2}, \quad (10)$$

The subscripts  $i = 1$  and  $2$  denote the two components.  $\sigma_{\text{TRK}}$  expresses the cross section in terms of the Thomas-Reiche-Kuhn sum rule. The Lorentzian function is applied before mapping the hidden nodes to the output cross sections, as shown in Fig. 1(c). Except for the Lorentzian shape, the known knowledge includes the empirical formula of the Lorentzian parameters:

$$\begin{aligned} E_1 &= H_1 A^{-1/3} - H_2 \beta_2, \\ E_2 &= H_1 A^{-1/3} + H_2 \beta_2, \\ \Gamma_1 &= H_3 - H_4 \beta_2, \\ \Gamma_2 &= H_3 + H_4 \beta_2, \\ s_1 &= H_5 - H_6 \beta_2, \\ s_2 &= H_5 + H_6 \beta_2, \end{aligned} \quad (11)$$

where  $H_i$  ( $i = 1, \dots, 6$ ) are empirical parameters. These empirical formulas are also considered in the BNN. The algorithm then becomes a Lorentzian function-based BNN (LBNN). As shown in Fig. 1(c), the (black) solid lines show that all inputs nodes are used to calculate the empirical parameters  $H_1, H_2, H_3, H_4, H_5$ , and  $H_6$ . Then, the empirical parameters, together with the input nodes  $A$  and  $\beta_2$ , are used to calculate the Lorentzian parameters  $E_1, E_2, \Gamma_1, \Gamma_2, s_1$ , and  $s_2$  according to Eq. (11). As can be observed, the Lorentzian parameters depend not only on  $A$  and  $\beta_2$  but also on other input nodes. However, the dependence on  $A$  and  $\beta_2$  is known and physical, as shown in Eq. (11), whereas the dependence on all the input nodes is unknown and numerical. These were fitted during the training process. We use two types of lines, solid and dashed, to distinguish between the physical and numerical dependences. Energy is one of the input nodes, but is not used to calculate the hidden layer. This is because the dependence of the energy on the output (cross section) is known as the Lorentzian shape. In the BNN-ZAE and BNN-OPT models, the dependence on energy of the output (cross section) is a black box. In the LBNN model, some dependencies on the inputs of the outputs are known. We say that parts of the black box are open.

### III. RESULTS & DISCUSSIONS

There are three types of BNN in this work: BNN-ZAE, BNN-OPT, and LBNN. In the latter two, physics guidance is provided by improving the input nodes and considering the

Lorentzian shape. In contrast, BNN-ZAE is a numerical algorithm without any physical improvement. In the following section, we evaluate these three models by comparing their predictions. The root-mean-square (*RMS*) deviations between the predictions and data were calculated as

$$RMS = \sqrt{\frac{1}{N_d} \sum_{n=1}^{N_d} \left( \log \sigma_p^{(n)} - \log \sigma_d^{(n)} \right)^2}, \quad (12)$$

where  $\sigma_p$  is the predicted cross section and  $\sigma_d$  is the corresponding data. Log-scaling was used because the values of  $\sigma_d$  are in the region from  $10^{-3}$  to  $10$  b, which is across four orders of magnitude.

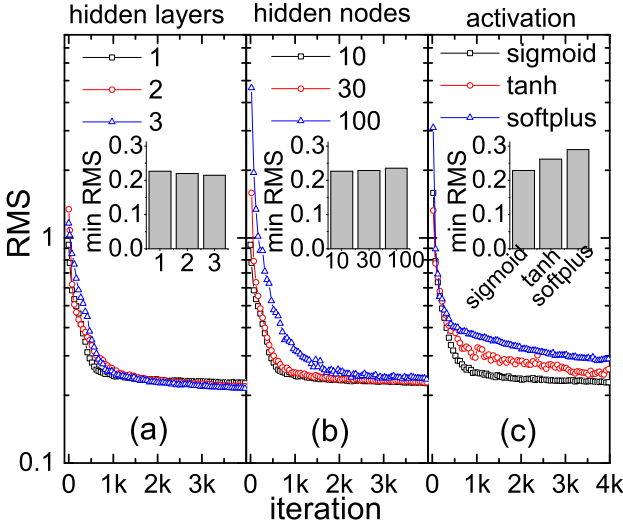


Fig. 2. (Color online) Root-mean-square (*RMS*) deviations between data and the BNN-ZAE predictions as a function of the iteration step. The insets show the *RMS* deviations at the iteration step 4000. (a) Comparison of the *RMS* deviations for 1, 2, and 3 hidden layers. (b) Comparison of the *RMS* deviations for 1 hidden layer with 10, 30, and 100 hidden nodes. (c) Comparison of the *RMS* deviations for 1 hidden layer with 30 hidden nodes but different activation functions.

The *RMS* deviations as a function of the iteration step were applied to test the convergence of the predictions. The cases for the BNN-ZAE predictions with one, two, and three hidden layers and 10 nodes for each hidden layer are shown in Fig. 2(a). From the figure, we can observe that all the *RMS* deviations converge quickly in 1000 iteration steps. The final *RMS* deviations were 0.226, 0.219, and 0.214 for the cases of one, two, and three hidden layers, respectively. This indicates that more hidden layers help reproduce the training data; however, the effect is weak. The effect of the number of hidden nodes was also tested by setting one hidden layer to 10, 30, and 100 nodes. The *RMS* deviations are compared in Fig. 2(b). It was shown that the rate of convergence was slower for a larger number of hidden nodes. The iteration steps required to decrease the *RMS* deviations to a value less than 0.25 are 1000 for the cases of 10 and 30 hidden nodes, but 2000 steps are needed for 100 hidden nodes. The *RMS*

deviations at 4000 iteration steps are 0.226, 0.229, and 0.235 for 10, 30, and 100 nodes, respectively. The *RMS* deviations are similar for 10 and 30 hidden nodes, but they are larger for 100 hidden nodes. The role of the activation function was tested by setting a hidden layer with 30 hidden nodes using the sigmoid, tanh, or softplus activation function, as shown in Eq. (2). The *RMS* deviations are compared in Fig. 2(c). The sigmoid activation function was shown to be the best. In the following calculations, three hidden layers with 30 hidden nodes for each layer and a sigmoid activation function were applied.

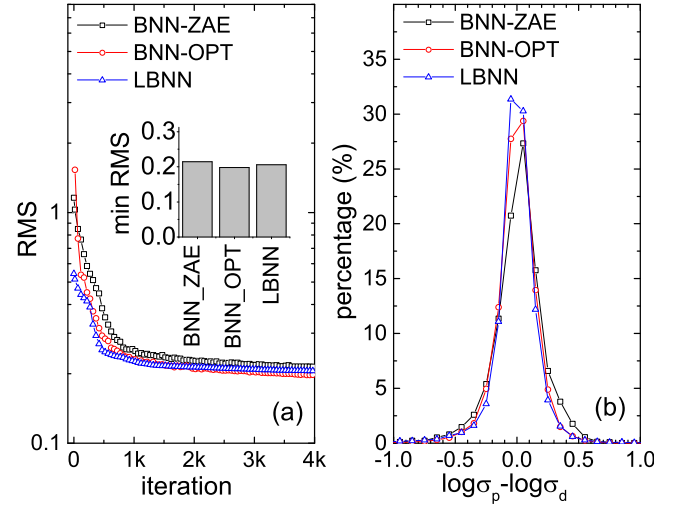


Fig. 3. (Color online) Comparison of the prediction errors for three types of Bayesian neural networks (as shown in Fig. 1). (a) The *RMS* deviations as a function of the iteration step. The inset shows the *RMS* deviations at the iteration step 4000. (b) Distributions of the differences between predictions and training data  $\log \sigma_p - \log \sigma_d$ .

The *RMS* deviations as a function of the iteration step for the three types of BNNs are shown in Fig. 3(a). It is shown that the rates of convergence are similar for all the cases, and 4000 iteration steps are sufficient for the calculations. The *RMS* deviations at 4000 iteration steps are 0.198 for the BNN-OPT model and 0.206 for the LBNN model, which are smaller than that of 0.214 for BNN-ZAE. Figure 2(a) and (b) shows that a change in the numerical parameters (number of hidden layers and hidden nodes) does not improve the neural network. Here, it is indicated that the key to improving the neural network for physics is the consideration of the known knowledge of observables. For instance, several effects of the photoneutron reaction, such as isospin dependence and shape effect, have been found [3, 39, 40], which indicate that the cross sections of the photoneutron reaction depend on the ground-state properties of the nuclei. Another known knowledge of photoneutron cross sections is their Lorentzian shape. The data of the ground-state properties are applied to the input layer of the BNN-OPT model, whereas the Lorentzian shape is considered in the LBNN model. These two aspects are responsible for the smaller *RMS* deviations compared to those provided by the BNN-ZAE algorithm.



To compare the errors of the predictions by the three neural networks, the distributions of the differences between the predictions and training data  $\log \sigma_p - \log \sigma_d$  are shown in Fig. 3(b). A value 1 of  $\log \sigma_p - \log \sigma_d$  means that the predicted cross section is 10 times larger than the experimental data, whereas a value of -1 indicates that it is 10 times smaller. These cases hardly ever occur. The sample size for  $abs(\log \sigma_p - \log \sigma_d) > 1$  is less than 0.1% of the total sample size. For most samples, the values of  $\log \sigma_p - \log \sigma_d$  are in the region from -0.1 to 0.1, which means that the prediction agrees with the data in 1.26 times. Specifically, these samples comprise 48.0%, 57.1%, and 61.6% of BNN-ZAE, BNN-OPT, and LBNN samples, respectively. The percentage of LBNN was the largest. Furthermore, the distribution of  $\log \sigma_p - \log \sigma_d$  provided by the LBNN was more symmetrical than that of the others. Considering these aspects, the LBNN model is superior to the BNN-OPT model.

We further evaluated the three types of BNNs by comparing their predictions of photoneutron yield cross sections for spherical nuclei  $^{92}\text{Zr}$ ,  $^{112}\text{Sn}$ , and  $^{206}\text{Pb}$ . The cross sections as a function of the incident energy are shown in Fig. 4. In general, for spherical nuclei, only one main component of the Lorentzian shape is displayed in the excitation function of the photoneutron reaction. The experimental cross sections as a function of the incident energy  $\varepsilon$  for nuclei  $^{92}\text{Zr}$ ,  $^{112}\text{Sn}$ , and  $^{206}\text{Pb}$  display this situation. The position and value of the Lorentzian peak were reproduced well by the LBNN model. In contrast, the peak position for nucleus  $^{92}\text{Zr}$  and the peak value for nucleus  $^{112}\text{Sn}$  were underestimated by both the BNN-ZAE and BNN-OPT models.

The data were abundant for these three nuclei. Some experimental errors, including statistical and systematic errors, are shown as error bars in the figure. The error of the data was applied to train the BNN. More specifically, the data were resampled with a number inversely proportional to the experimental errors. This means that the data with small errors are resampled with a large number for training, while those with large errors are applied only a few times. After 4000 iteration steps, 100 samples are applied to calculate the standard deviations and uncertainties of the predictions, which are shown as shadows in the figure. In the energy region with data, the uncertainties of the predictions by the BNN-ZAE and BNN-OPT models are small, but there should be large uncertainties in the extrapolations. Because the Lorentzian function is applied to the neural network, the predictions and their uncertainties are both constrained by the Lorentzian shape. The uncertainties of the predictions by the LBNN model were the same for interpolations and extrapolations in logarithmic coordinates. Furthermore, the uncertainties are small because the data for these three nuclei are abundant. The uncertainties shown in Fig. 4 originate from the Monte Carlo techniques. The Lorentzian function applied in the LBNN model is only an approximate expression for the photoneutron yield cross sections. A threshold exists for the photoneutron yield cross sections at  $\varepsilon = S_n$ , where  $S_n$  is the neutron separation energy. The Lorentzian shape is the most important known knowledge of photoneutron yield cross sections, but it does not consider the threshold. The experimental data near the

threshold were against this formula. Thus, predictions by the LBNN model below the threshold are meaningless.

For the axially deformed nucleus, the photoneutron yield cross sections as a function of incident energy display two main Lorentzian shapes. The difference between the two Lorentzian peaks has been found to be positively associated with the deformation parameter  $\beta_2$  using the time-dependent Hartree-Fock model [41]. The data for deformed nuclei  $^{31}\text{P}$ ,  $^{75}\text{As}$ , and  $^{165}\text{Ho}$  are shown in Fig. 5. The quadrupole deformation parameters  $\beta_2$  are -0.22, -0.24, and 0.28, respectively. The data clearly show two peaks for  $^{165}\text{Ho}$ , but faintly for  $^{75}\text{As}$ . The abundant data but large errors for  $^{31}\text{P}$  make it difficult to distinguish the two peaks. However, the wide peak does not contradict the two main Lorentzian shapes because they may be too close to be distinguished.

The curves and shadows show the predictions with confidence intervals using neural networks. The BNN-ZAE model reproduces the overall trend for the data of  $^{31}\text{P}$  and  $^{75}\text{As}$  and slightly overestimates the cross sections of  $^{31}\text{P}$  in the most energy region. However, it neglects the two obvious peaks for  $^{165}\text{Ho}$  and predicts the cross sections with only one peak. At energy  $\varepsilon = 14$  MeV, where the data are approximately 0.25 b, the BNN-ZAE model grossly overestimates the cross section. This process is illustrated in Fig. 3. The BNN-OPT model provides a smaller RMS deviation than the BNN-ZAE model. This point is reiterated by comparing the predictions of the BNN-OPT and BNN-ZAE models. However, the confidence intervals provided by the BNN-OPT model were wider than those provided by the BNN-ZAE model. When extrapolating the cross section to the low-energy region,  $\varepsilon < 10$  MeV for  $^{31}\text{P}$  and  $^{75}\text{As}$ , overfitting by the BNN-OPT model is shown. The confidence intervals show an increasing cross section with decreasing energy, which was not observed in the experiment.

The LBNN model reproduced the data better than the other two models did. Two Lorentzian shapes related to the quadrupole deformation parameters  $\beta_2$  were considered in the LBNN model, as shown in Fig. 1(c), hence the predictions show two peaks for deformed nuclei. It cannot be denied that the LBNN model underestimates the cross section of  $^{31}\text{P}$  at  $\varepsilon = 21.5$  MeV, where the calculations show a valley between two Lorentzian peaks, but the data display a weak peak. The weak peak may reveal a substructure beyond the main Lorentzian shapes, which can also be observed in the data of  $^{206}\text{Pb}$  near  $\varepsilon = 25$  MeV. However, its origin has not been explained by the physics-motivated model, and hence, it has not been considered in the neural network. The improvement from BNN-OPT to LBNN supports the idea that the substructure beyond the main Lorentzian shapes can be considered in the neural network after its properties have been revealed.

#### IV. CONCLUSION

In conclusion, the photoneutron yield cross sections as a function of the charge number  $Z$ , mass number  $A$ , and incident energy  $\varepsilon$  were studied using the BNN, and the model is abbreviated as BNN-ZAE. The numerical parameters of the

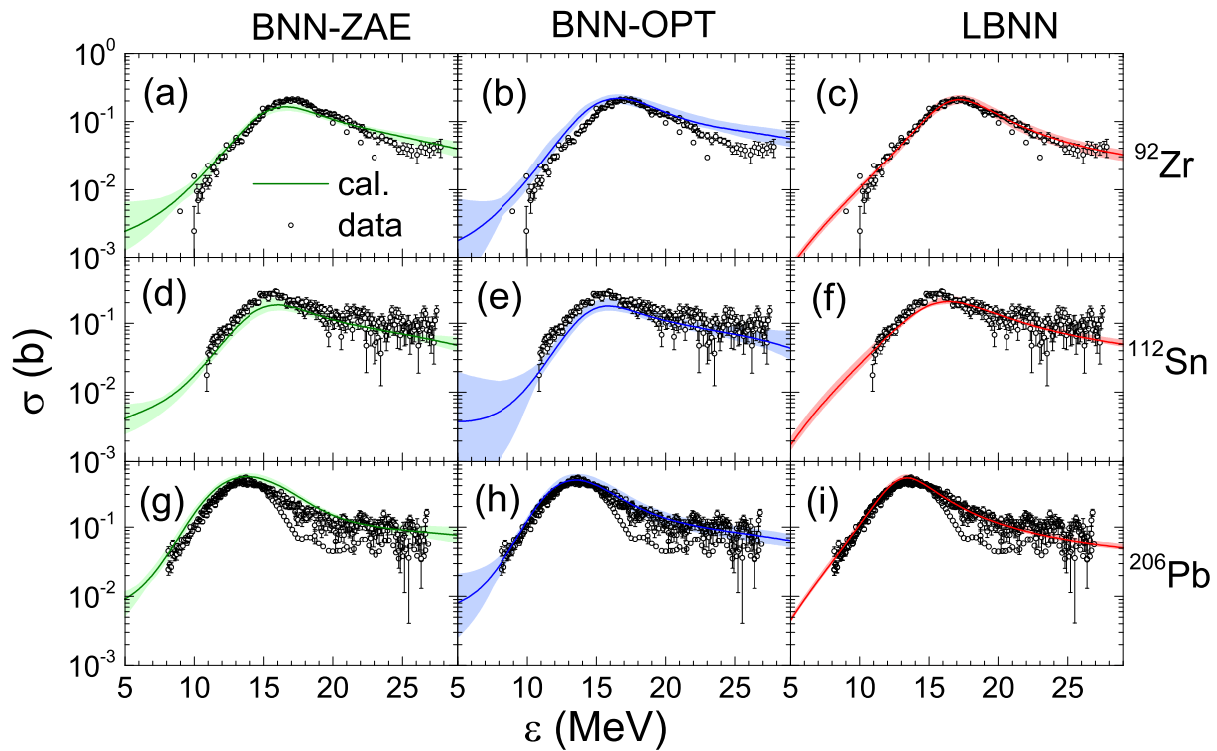


Fig. 4. (Color online) Photoneutron yield cross sections for spherical nuclei  $^{92}\text{Zr}$ ,  $^{112}\text{Sn}$ , and  $^{206}\text{Pb}$  as a function of the incident energy. The circles with error bars show the experimental data taken from the EXFOR database. The curves with shadows show the predictions of the three types of Bayesian neural networks (as shown in Fig. 1).

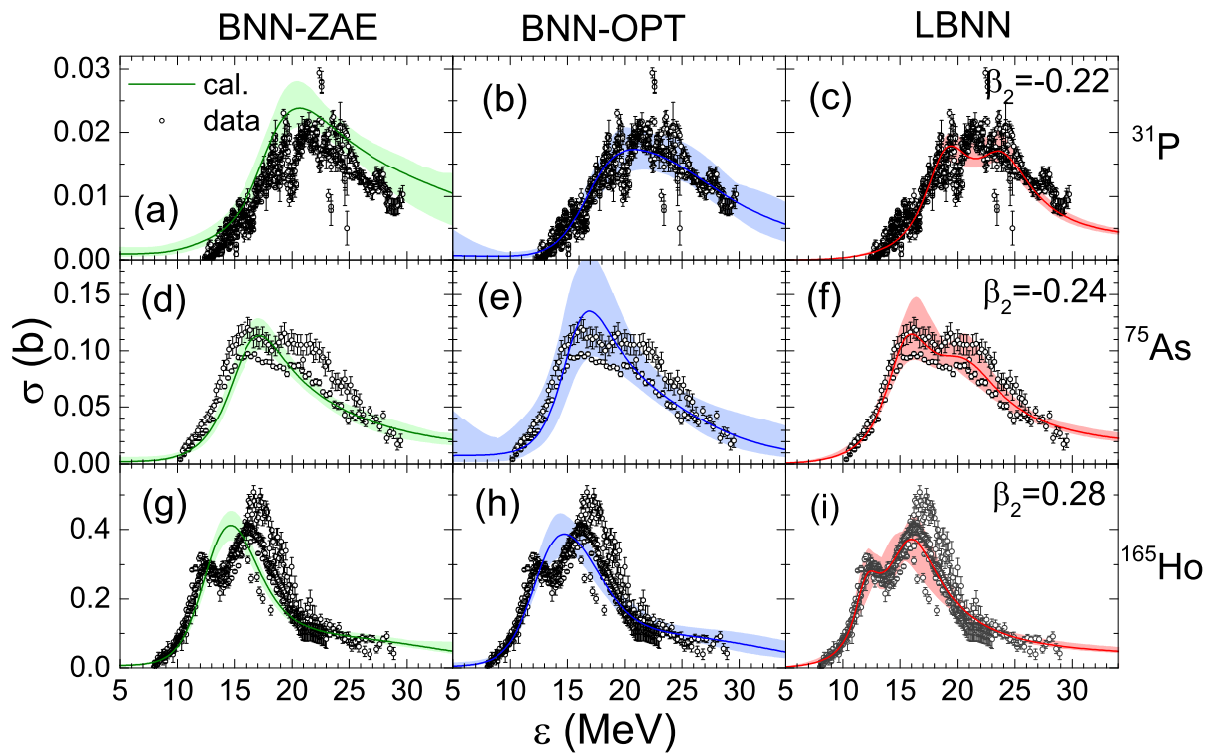


Fig. 5. (Color online) Same as Fig. 4 but for the deformed nuclei  $^{31}\text{P}$ ,  $^{75}\text{As}$ , and  $^{165}\text{Ho}$ . The quadrupole deformation parameters  $\beta_2$  of those nuclei are shown.

neural network were varied to test the model. The influence of the activation function on the prediction was determined. The

sigmoid activation function was the best for realizing nonlinearity between the input and output, and hence it provided the smallest deviations between predictions and data. However, the predictions of the BNN-ZAE model could not be improved by increasing the number of hidden layers from 1 to 3 and the number of hidden nodes from 10 to 100.

In the method proposed in Ref. [38], physics guidance is provided in BNNs from the input layer. Several effects of the photoneutron reaction, such as the isospin dependence and shape effect, have been observed [3, 39, 40], which indicate that the cross sections of the photoneutron reaction depend on the ground-state properties of the nuclei. Based on this knowledge, the optimal ground-state properties were selected as input neurons, resulting in the BNN-OPT model. It was shown that the deviations between the predictions of the BNN-OPT model and the data were smaller than those of the BNN-ZAE model.

The BNN was further improved by the Lorentzian shape of the photoneutron yield cross sections. The Lorentzian function was applied to map the hidden nodes to the output cross sections, and the empirical formula of the Lorentzian parameters was applied to link the input nodes to the output cross sections. This new algorithm is called Lorentzian function-based BNN (LBNN). We evaluated the BNN-ZAE, BNN-

OPT, and LBNN models by comparing their predictions of the photoneutron yield cross sections for the spherical nuclei  $^{92}\text{Zr}$ ,  $^{112}\text{Sn}$ , and  $^{206}\text{Pb}$ , as well as the deformed nuclei  $^{31}\text{P}$ ,  $^{75}\text{As}$ , and  $^{165}\text{Ho}$ . Generally, for spherical nuclei, only one main component of the Lorentzian shape exists. All three models reproduced the main trend of the data, but the predictions of the LBNN model were the best. For an axially deformed nucleus, the photoneutron yield cross sections displayed two main Lorentzian shapes. Only the LBNN model reproduced two peaks of the cross sections in the deformed nuclei  $^{31}\text{P}$ ,  $^{75}\text{As}$ , and  $^{165}\text{Ho}$ . This is because two Lorentzian shapes related to quadrupole deformation parameters  $\beta_2$  are considered in the LBNN model.

## AUTHOR CONTRIBUTIONS

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Yong-Yi Li, Fan Zhang and Jun Su. The first draft of the manuscript was written by Yong-Yi Li and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

- 
- [1] Z. Niu, H. Liang, Nuclear mass predictions based on bayesian neural network approach with pairing and shell effects. *Phys. Lett. B* **778**, 48–53 (2018). Doi: [10.1016/j.physletb.2018.01.002](https://doi.org/10.1016/j.physletb.2018.01.002)
  - [2] Z.A. Wang, J. Pei, Y. Liu, et al., Bayesian evaluation of incomplete fission yields. *Phys. review letters* **123**, 122501 (2019). Doi: [10.1103/PhysRevLett.123.122501](https://doi.org/10.1103/PhysRevLett.123.122501)
  - [3] R. Wang, Z. Zhang, L.W. Chen, et al., Constraining the in-medium nucleon-nucleon cross section from the width of nuclear giant dipole resonance. *Phys. Lett. B* **807**, 135532 (2020). Doi: [10.1016/j.physletb.2020.135532](https://doi.org/10.1016/j.physletb.2020.135532)
  - [4] X.C. Ming, H.F. Zhang, R.R. Xu, et al., Nuclear mass based on the multi-task learning neural network method. *Nucl. Sci. Tech.* **33**, 48 (2022). Doi: [10.1007/s41365-022-01031-z](https://doi.org/10.1007/s41365-022-01031-z)
  - [5] Z.P. Gao, Y.J. Wang, H.L. Lü, et al., Machine learning the nuclear mass. *Nucl. Sci. Tech.* **32**, 109 (2021). Doi: [10.1007/s41365-021-00956-1](https://doi.org/10.1007/s41365-021-00956-1)
  - [6] C.W. Ma, X.B. Wei, X.X. Chen, et al., Precise machine learning models for fragment production in projectile fragmentation reactions using bayesian neural networks. *Chin. Phys. C* **46**, 074104 (2022). Doi: [10.1088/1674-1137/ac5efb](https://doi.org/10.1088/1674-1137/ac5efb)
  - [7] K. Gernoth, J. Clark, J. Prater, et al., Neural network models of nuclear systematics. *Phys. Lett. B* **300**, 1–7 (1993). Doi: [10.1016/0370-2693\(93\)90738-4](https://doi.org/10.1016/0370-2693(93)90738-4)
  - [8] S. Athanassopoulos, E. Mavrommatis, K. Gernoth, et al., Nuclear mass systematics using neural networks. *Nucl. Phys. A* **743**, 222–235 (2004). Doi: [10.1016/j.nuclphysa.2004.08.006](https://doi.org/10.1016/j.nuclphysa.2004.08.006)
  - [9] D. Benzaid, S. Bentradi, A. Kerraci, et al., Bethe–weizsäcker semiempirical mass formula coefficients 2019 update based on ame2016. *Nucl. Sci. Tech.* **31**, 9 (2020). Doi: [10.1007/s41365-019-0718-8](https://doi.org/10.1007/s41365-019-0718-8)
  - [10] N. Costiris, E. Mavrommatis, K.A. Gernoth, et al., Decoding  $\beta$ -decay systematics: A global statistical model for  $\beta$ -half-lives. *Phys. Rev. C* **80**, 044332 (2009). Doi: [10.1103/PhysRevC.80.044332](https://doi.org/10.1103/PhysRevC.80.044332)
  - [11] T. Bayram, S. Akkoyun, S.O. Kara, A study on ground-state energies of nuclei by using neural networks. *Annals Nucl. Energy* **63**, 172–175 (2014). Doi: [10.1016/j.anucene.2013.07.039](https://doi.org/10.1016/j.anucene.2013.07.039)
  - [12] Z. Niu, H. Liang, B. Sun, et al., Predictions of nuclear  $\beta$ -decay half-lives with machine learning and their impact on r-process nucleosynthesis. *Phys. Rev. C* **99**, 064307 (2019). Doi: [10.1103/PhysRevC.99.064307](https://doi.org/10.1103/PhysRevC.99.064307)
  - [13] W. Jiang, G. Hagen, T. Papenbrock, Extrapolation of nuclear structure observables with artificial neural networks. *Phys. Rev. C* **100**, 054326 (2019). Doi: [10.1103/PhysRevC.100.054326](https://doi.org/10.1103/PhysRevC.100.054326)
  - [14] X.X. Dong, R. An, J.X. Lu, et al., Novel bayesian neural network based approach for nuclear charge radii. *Phys. Rev. C* **105**, 014308 (2022). Doi: [10.1103/PhysRevC.105.014308](https://doi.org/10.1103/PhysRevC.105.014308)
  - [15] D. Wu, C. Bai, H. Sagawa, et al., Calculation of nuclear charge radii with a trained feed-forward neural network. *Phys. Rev. C* **102**, 054323 (2020). Doi: [10.1103/PhysRevC.102.054323](https://doi.org/10.1103/PhysRevC.102.054323)
  - [16] X. Zhang, X. Liu, Y. Huang, et al., Determining impact parameters of heavy-ion collisions at low-intermediate incident energies using deep learning with convolutional neural networks. *Phys. Rev. C* **105**, 034611 (2022). Doi: [10.1103/PhysRevC.105.034611](https://doi.org/10.1103/PhysRevC.105.034611)
  - [17] X.Z. Li, Q.X. Zhang, H.Y. Tan, et al., Fast nuclide identification based on a sequential bayesian method. *Nucl. Sci. Tech.* **32**, 143 (2021). Doi: [10.1007/s41365-021-00982-z](https://doi.org/10.1007/s41365-021-00982-z)
  - [18] H.R. Liu, Y.X. Cheng, Z. Zuo, et al., Discrimination of neutrons and gamma rays in plastic scintillator based on pulse-coupled neural network. *Nucl. Sci. Tech.* **32**, 82 (2021). Doi: [10.1007/s41365-021-00915-w](https://doi.org/10.1007/s41365-021-00915-w)
  - [19] T.Y. Huang, Z.G. Li, K. Wang, et al., Hybrid windowed networks for on-the-fly doppler broadening in rmc code. *Nucl. Sci.*

- Tech. **32**, 62 (2021). Doi: [10.1007/s41365-021-00901-2](https://doi.org/10.1007/s41365-021-00901-2)
- [20] K. Chen, L.B. Zhang, J.S. Liu, et al., Robust restoration of low-dose cerebral perfusion ct images using ncs-unet. Nucl. Sci. Tech. **33**, 30 (2022). Doi: [10.1007/s41365-022-01014-0](https://doi.org/10.1007/s41365-022-01014-0)
- [21] Y.J. Ma, Y. Ren, P. Feng, et al., Sinogram denoising via attention residual dense convolutional neural network for low-dose computed tomography. Nucl. Sci. Tech. **32**, 41 (2021). Doi: [10.1007/s41365-021-00874-2](https://doi.org/10.1007/s41365-021-00874-2)
- [22] R. Utama, J. Piekarewicz, H. Prosper, Nuclear mass predictions for the crustal composition of neutron stars: A bayesian neural network approach. Phys. Rev. C **93**, 014311 (2016). Doi: [10.1103/PhysRevC.93.014311](https://doi.org/10.1103/PhysRevC.93.014311)
- [23] R. Utama, W.C. Chen, J. Piekarewicz, Nuclear charge radii: density functional theory meets bayesian neural networks. J. Phys. G: Nucl. Part. Phys. **43**, 114002 (2016). Doi: [10.1088/0954-3899/43/11/114002](https://doi.org/10.1088/0954-3899/43/11/114002)
- [24] L. Neufcourt, Y. Cao, W. Nazarewicz, et al., Bayesian approach to model-based extrapolation of nuclear observables. Phys. Rev. C **98**, 034318 (2018). Doi: [10.1103/PhysRevC.98.034318](https://doi.org/10.1103/PhysRevC.98.034318)
- [25] C.W. Ma, D. Peng, H.L. Wei, et al., Isotopic cross-sections in proton induced spallation reactions based on the bayesian neural network method. Chin. Phys. C **44**, 014104 (2020). Doi: [10.1088/1674-1137/44/1/014104](https://doi.org/10.1088/1674-1137/44/1/014104)
- [26] J. Bai, Z. Niu, B. Sun, et al., The description of giant dipole resonance key parameters with multitask neural networks. Phys. Lett. B **815**, 136147 (2021). Doi: [10.1016/j.physletb.2021.136147](https://doi.org/10.1016/j.physletb.2021.136147)
- [27] X. Wang, L. Zhu, J. Su, Providing physics guidance in bayesian neural networks from the input layer: The case of giant dipole resonance predictions. Phys. Rev. C **104**, 034317 (2021). Doi: [10.1103/PhysRevC.104.034317](https://doi.org/10.1103/PhysRevC.104.034317)
- [28] G. Baldwin, G. Klaiber, Photo-fission in heavy elements. Phys. Rev. **71**, 3 (1947). Doi: [10.1103/PhysRev.71.3](https://doi.org/10.1103/PhysRev.71.3)
- [29] M.T. Jin, S.Y. Xu, G.M. Yang, et al., Yield of long-lived fission product transmutation using proton-, deuteron-, and alpha particle-induced spallation. Nucl. Sci. Tech. **32**, 96 (2021). Doi: [10.1007/s41365-021-00933-8](https://doi.org/10.1007/s41365-021-00933-8)
- [30] V. Plujko, O. Gorbachenko, R. Capote, et al., Giant dipole resonance parameters of ground-state photoabsorption: Experimental values with uncertainties. At. Data Nucl. Data Tables **123**, 1–85 (2018). Doi: [10.1016/j.adt.2018.03.002](https://doi.org/10.1016/j.adt.2018.03.002)
- [31] T. Kawano, Y. Cho, P. Dimitriou, et al., Iaea photonuclear data library 2019. Nucl. Data Sheets **163**, 109–162 (2020). Doi: [10.1016/j.nds.2019.12.002](https://doi.org/10.1016/j.nds.2019.12.002)
- [32] D. Savran, T. Aumann, A. Zilges, Experimental studies of the pygmy dipole resonance. Prog. Part. Nucl. Phys. **70**, 210–245 (2013). Doi: [10.1016/j.ppnp.2013.02.003](https://doi.org/10.1016/j.ppnp.2013.02.003)
- [33] A. Bracco, E. Lanza, A. Tamii, Isoscalar and isovector dipole excitations: Nuclear properties from low-lying states and from the isovector giant dipole resonance. Prog. Part. Nucl. Phys. **106**, 360–433 (2019). Doi: [10.1016/j.ppnp.2019.02.001](https://doi.org/10.1016/j.ppnp.2019.02.001)
- [34] V. Semkova, N. Otuka, M. Mikhailiukova, et al., in *EPJ Web of Conferences*, Exfor—a global experimental nuclear reaction data repository: Status and new developments. Vol. 146, EDP Sciences, 2017, p. 07003, doi: [10.1051/epjconf/201714607003](https://doi.org/10.1051/epjconf/201714607003)
- [35] J.Y. Tang, Q. An, J.B. Bai, et al., Back-n white neutron source at cns and its applications. Nucl. Sci. Tech. **32**, 11 (2021). Doi: [10.1007/s41365-021-00846-6](https://doi.org/10.1007/s41365-021-00846-6)
- [36] Y.T. Li, W.P. Lin, B.S. Gao, et al., Development of a low-background neutron detector array. Nucl. Sci. Tech. **33**, 41 (2022). Doi: [10.1007/s41365-022-01030-0](https://doi.org/10.1007/s41365-022-01030-0)
- [37] D.J. MacKay, A practical bayesian framework for backpropagation networks. Neural computation **4**, 448–472 (1992). Doi: [10.1162/neco.1992.4.3.448](https://doi.org/10.1162/neco.1992.4.3.448)
- [38] X. Wang, L. Zhu, J. Su, Providing physics guidance in bayesian neural networks from the input layer: The case of giant dipole resonance predictions. Phys. Rev. C **104**, 034317 (2021). Doi: [10.1103/PhysRevC.104.034317](https://doi.org/10.1103/PhysRevC.104.034317)
- [39] D. Pandit, S. Bhattacharya, D. Mondal, et al., Role of fluctuations in a thermal phase transition in a nucleus probed via the giant dipole resonance. Phys. Rev. C **99**, 024315 (2019). Doi: [10.1103/PhysRevC.99.024315](https://doi.org/10.1103/PhysRevC.99.024315)
- [40] J. Su, Constraining symmetry energy at subnormal density by isovector giant dipole resonances of spherical nuclei. Chin. Phys. C **43**, 064109 (2019). Doi: [10.1088/1674-1137/43/6/064109](https://doi.org/10.1088/1674-1137/43/6/064109)
- [41] A.A.B. Mennana, Y.E. Bassem, M. Oulne, Giant dipole resonance and shape evolution in nd isotopes within tdhf method. Phys. Scripta **95**, 065301 (2020). Doi: [10.1088/1402-4896/ab73d8](https://doi.org/10.1088/1402-4896/ab73d8)